Language Acquisition

Cognitive Development, Week 11

What does language comprehension entail?

What does language comprehension entail?
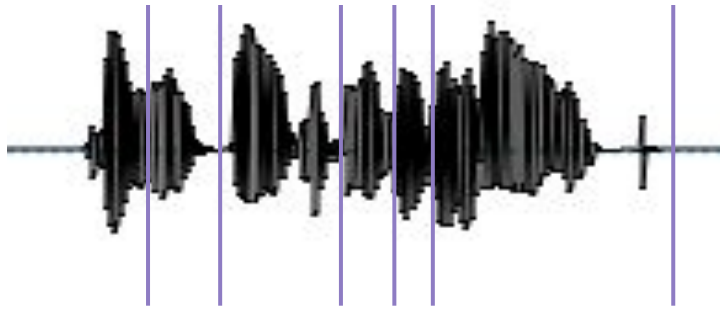
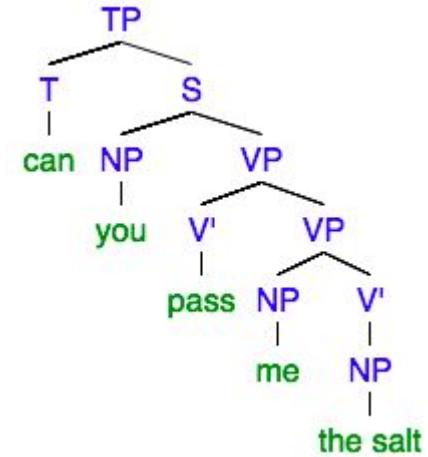can you pass me the salt

What are the words? (Phonology)

What does language comprehension entail?

can   you pass me the   salt
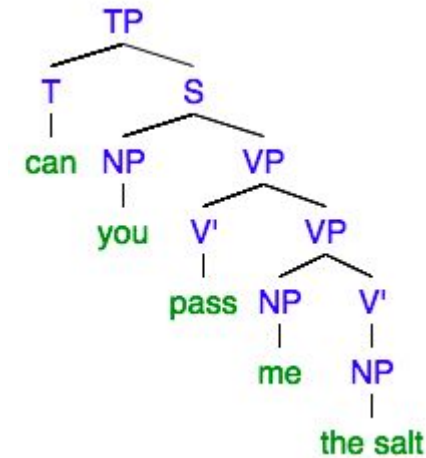
What are the words? (Phonology)

What is the structure? (Syntax)

What does language comprehension entail?

can   you pass me the   salt

What are the words? (Phonology)

What is the structure? (Syntax)

you: Agent (the person doing passing)
the salt: Patient (the thing being passed)
me: Recipient (the person receiving the salt)

Is it a question about whether I can pass the salt? Are they asking me to pass the salt?

What does it mean? (Semantics + Pragmatics)

How do children learn language?

Innate   Experience dependent
Domain-specific   Domain-general
Symbolic / Rule-based   No explicit rules

How you think about language acquisition depends on where you are on this spectrum

Outline

Arguments for innate, domain specific and modular mechanism

- Chomsky's argument (UG and LAD)
- Critical period
  - Genie
  - Language deprivation in deaf children
- Language creation (Nicaraguan Sign Language)
- Syntactic bootstrapping

Arguments for experience dependent and domain general mechanism

- Criticism of Chomsky's argument
- Statistical learning in word segmentation
- Role of bias in word learning
- Sensitivity to statistical regularities in syntactic structure
- (Morpho)-Syntax in connectionist networks

Arguments for an innate, domain-specific and rule based mechanism

Why is it symbolic and rule-based?

- People exhibit complex linguistic behaviour that cannot be explained without assuming that there is a "deep structure" to the sentence that is different from the "surface structure"
    - Aaroha danced
    - The vase broke
    - The vase was broken (by Aaroha)

    The first two have similar surface structure, but the last two have similar deep structure

- What is the deep structure? How do we get from deep structure to surface structure (or vice versa)?
    - Not possible to answer these questions without symbols and rules

Why is it symbolic and rule-based?

- Generative grammar (Chomsky, 1985): For any given sentence it must be able to assign a structure
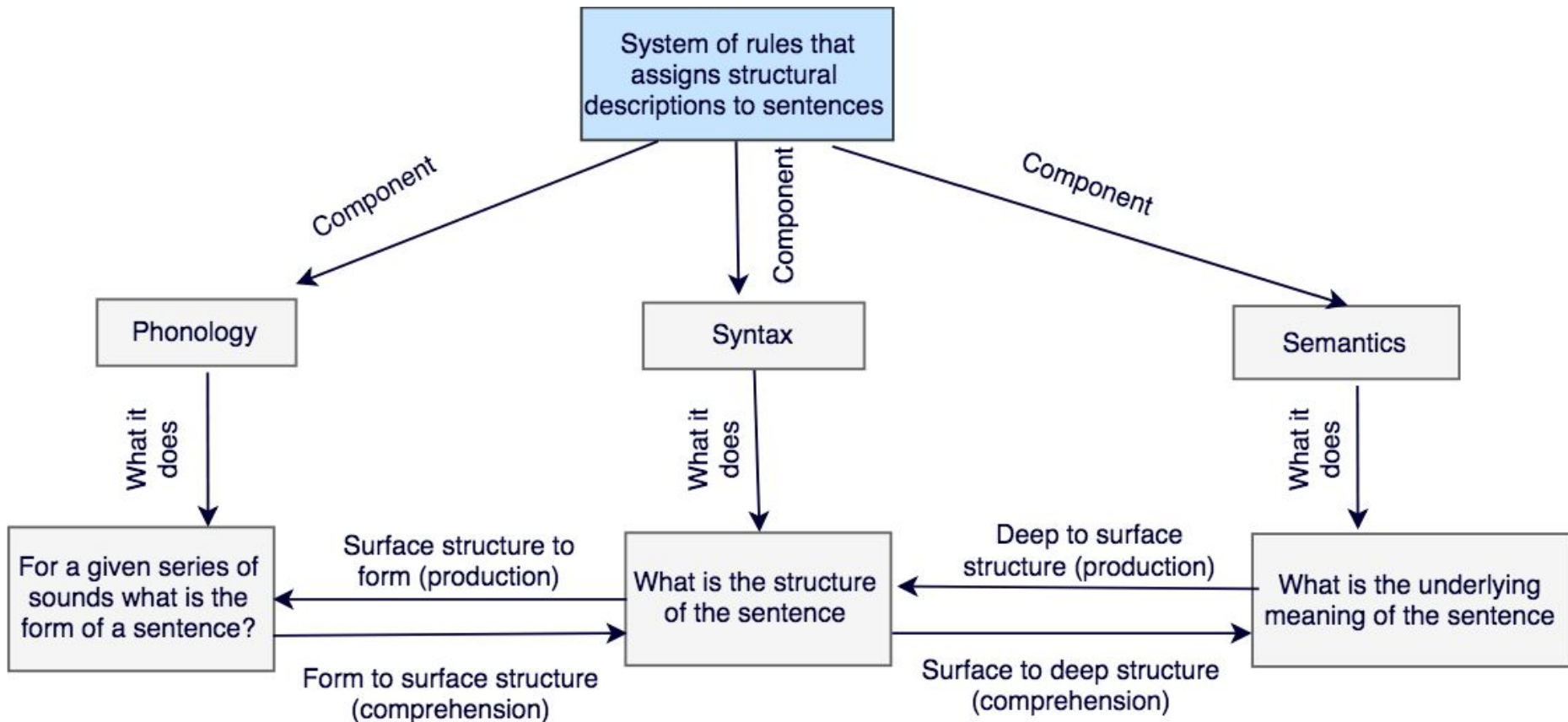
Why is it innate and domain-specific?

- There is a lot of linguistic diversity.
    - When will Aaroha come home?
    - Aaroha intiki eppudu vasthundi?
      Aaroha home when   come

- One solution: Come up with a different generative grammar for every language.

- Chomsky (Aspects): This is not "explanatorily adequate"
    - All children are born with the same architecture. How can the same architecture be capable of learning all these different grammars?

    - Could they learn from experience? Chomsky: No. Poverty of the stimulus argument

Why is it innate and domain-specific?

- Solution: Universal grammar. All languages have the same underlying principles but they vary in how they set the parameters.

- Language Acquisition Device (LAD)

  - Mechanism that is hard wired with these principles and allows children to set parameters based on input.

- These principles are very specific to language and do not make sense for other domains of cognition

  - E.g. Extended Projection Principle (EPP): All sentences must have subjects

Why is it innate and domain-specific?

- Solution: Universal grammar. All languages have the same underlying principles but they vary in how they set the parameters.

- Language Acquisition Device (LAD)
  - Mechanism that is hard wired with these principles and allows children to set parameters **based on input.**

- These principles are very specific to language and do not make sense for other domains of cognition
  - E.g. Extended Projection Principle (EPP): All sentences must have subjects

What happens when there is no input?

Critical period

Critical period

- "A critical period is a period of growth-in some cases only a few hours long-during which a specific kind of experience must occur or an ability or behavior will not develop" — LCC Chapter 1

- Lenneberg critical period hypothesis: Child needs to be exposed to language in the first few years, otherwise the child cannot learn language. Role of experience in nativist accounts.
  - How can we test this?

Genie

Genie

Thoughts?

Genie

Thoughts?

- The video said this case allows us to investigate "nature vs. nurture". Is it nature or nurture? How does it fit in with Chomsky's account?

- What does it mean to say she did not learn language?

- Any concerns about the conclusions we can draw from this case study?

Chomsky: https://www.youtube.com/watch?v=GmH6ffqKENA

Open class vs. closed class categories

- Open class: "Content words". Nouns, verbs, adjectives, adverbs
  - Change can be rapid
    - You got <u>crunk</u> at the party yesterday.
    - Did you <u>like</u> the photo I just uploaded?
  - Minimal effects of critical period
  - Other animals (e.g. Kanzi) successfully learned it

- Closed class: "Function words". Prepositions, pronouns, auxiliaries
  - Change is very slow
  - More drastic effects of critical period

Language deprivation in deaf children
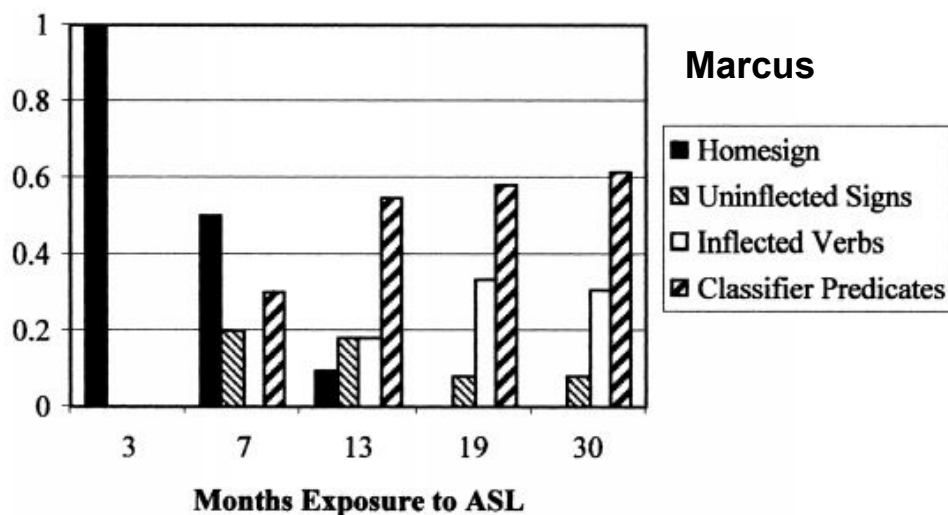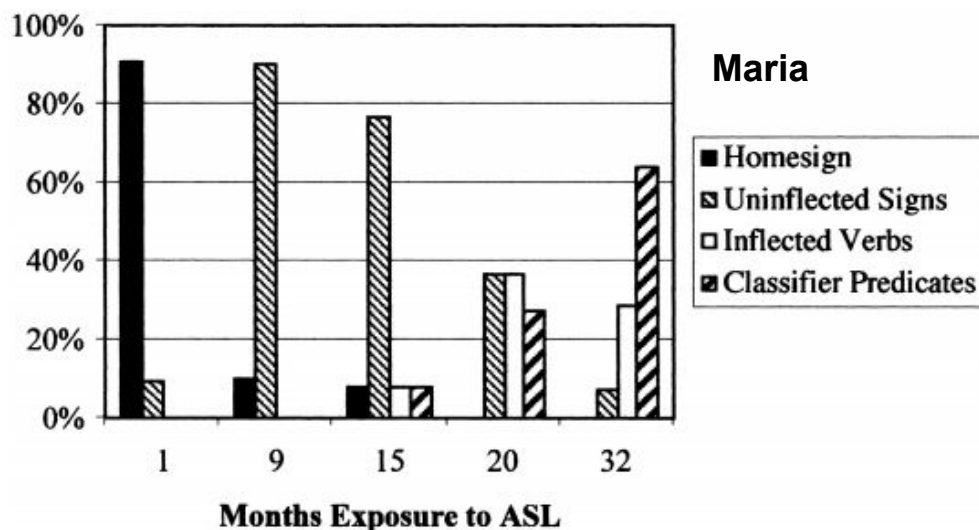
Language deprivation in deaf children

- Deaf children born to hearing parents do not get linguistic input if the parents do not use sign language.

- Sign language is not the same as using gestures to communicate
  - It is not purely iconic
  - It has morphosyntax (like verb agreement)
    - Children exposed to ASL from infancy acquire these at around the same age as children with spoken languages.
    - Make similar errors while learning: initially use uninflected forms, then overgeneralize.

Language deprivation in deaf children

Moford (2003)

- Maria (exposed to ASL at 13;7) and Marcus (exposed to ASL 12;1)

- Only communicated in homesign

- Enrolled in school where teachers used mostly ASL. And a little bit of signed English.

- Tasks

  - Production task: Narrate a story

  - Comprehension task: Watch a story being narrated and pick objects that corresponded to the question.

Language deprivation in deaf children

Maria

Marcus

Production task

There is individual variation, but both of them acquired morpho-syntax (as seen by their use of inflected verbs and classifier predicates)

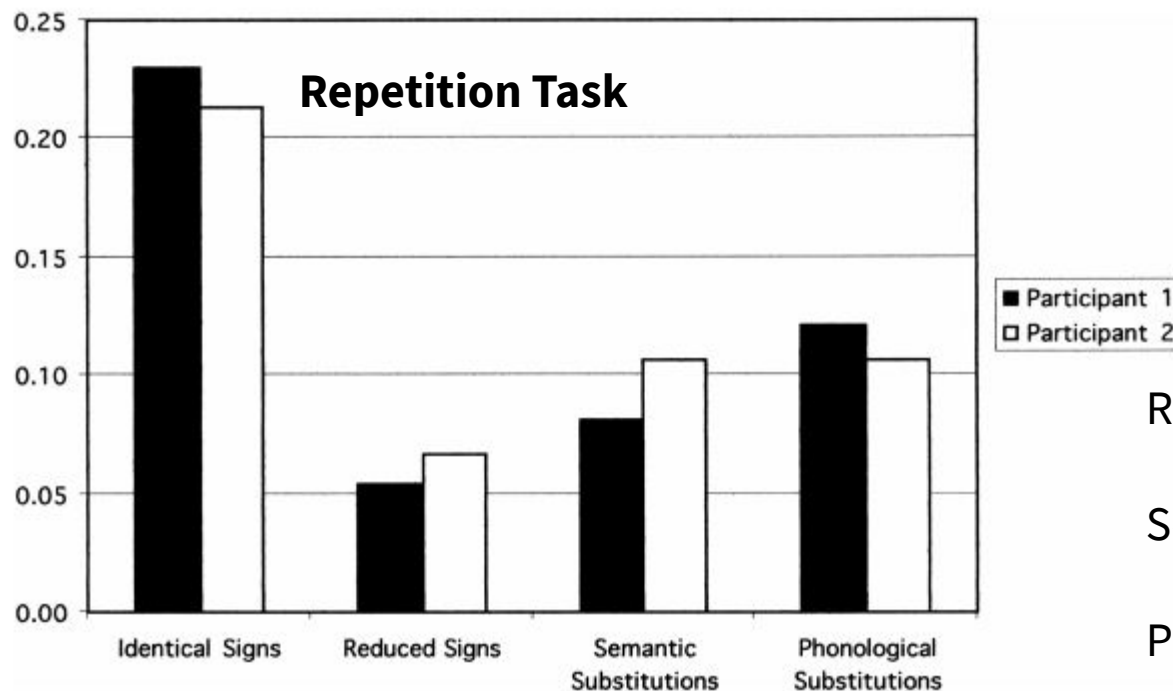# Language deprivation in deaf children

Comprehension task: (chance 25%)

- High processing load: correct picture 32% of the time
- Low processing load: Maria 63% and Marcus 100%

Language deprivation in deaf children

Comprehension task: (chance 25%)

- High processing load: correct picture 32% of the time
- Low processing load: Maria 63% and Marcus 100%

Reduced: No inflection

Semantic: Similar meaning

Phonological: Similar form
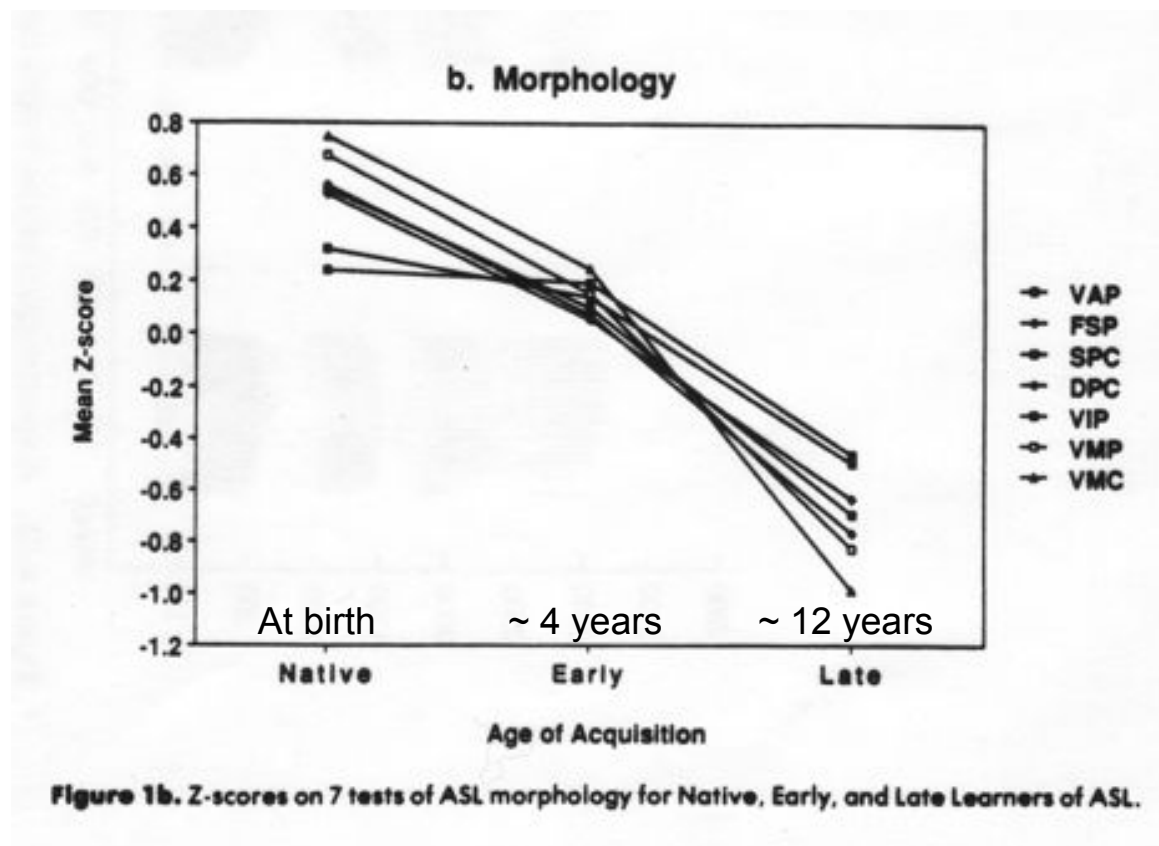
Language deprivation in deaf children

Conclusions

- Despite lack of exposure to language till adolescence, they were still able to acquire language — vocabulary + syntax

- Errors due to problems with processing? Competence vs performance.

Language deprivation in deaf children

Conclusions

- Despite lack of exposure to language till adolescence, they were still able to acquire language — vocabulary + syntax

- Errors due to problems with processing? Competence vs performance.

Critical period not so critical?

Language deprivation in deaf children

- Newport (1990): Compared language ability between 35-70 years. Minimum 30 years of exposure

b. Morphology

VAP:Agreement
VMC: Comprehension
VMP: Production

At birth   ~ 4 years   ~ 12 years

Native   Early   Late

Age of Acquisition

Figure 1b. Z-scores on 7 tests of ASL morphology for Native, Early, and Late Learners of ASL.
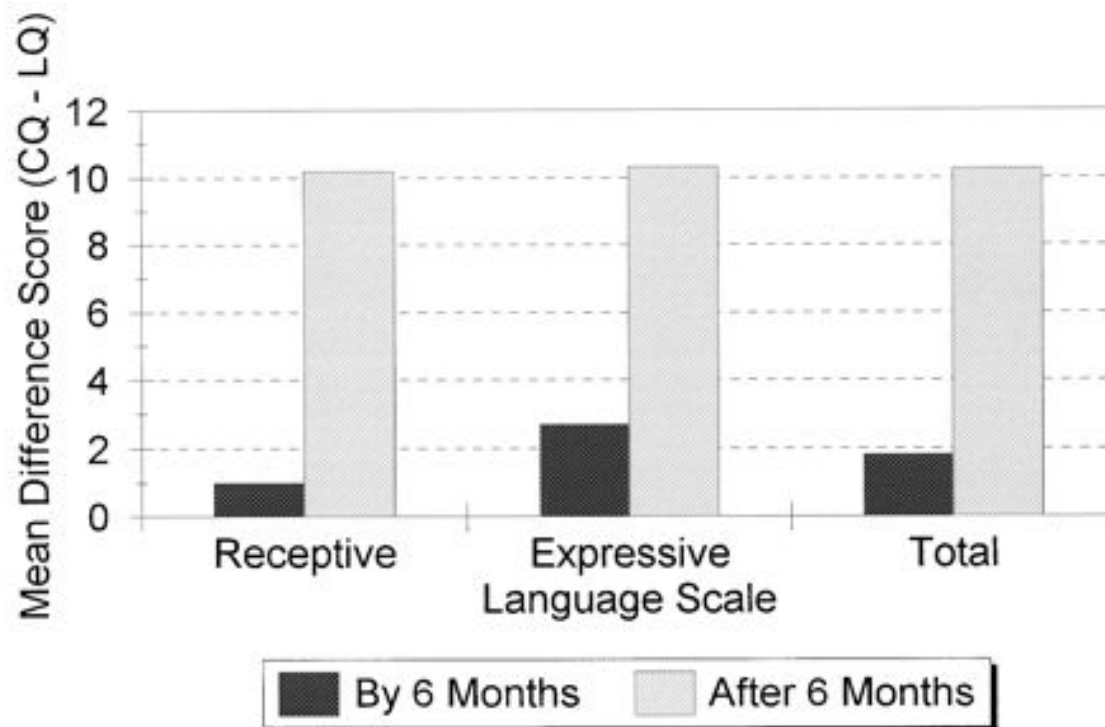
Language deprivation in deaf children

- Yoshinaga-Itano et al (1998): Compared language ability of children between 1 and 3 years of age.
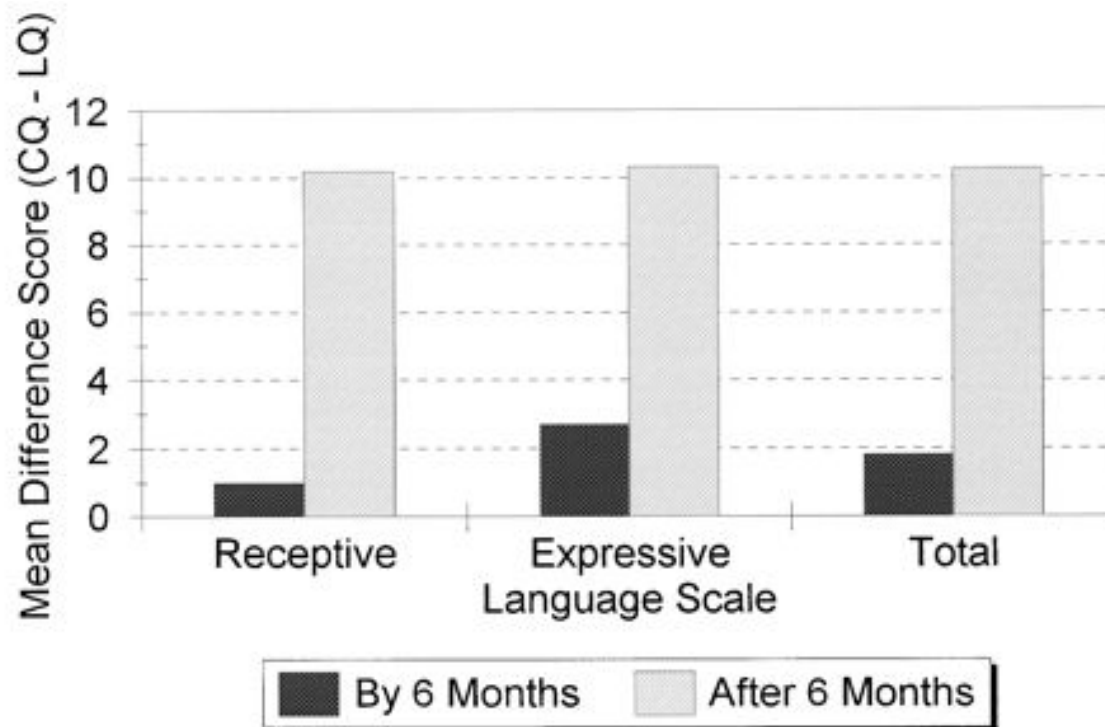
Language deprivation in deaf children

- Yoshinaga-Itano et al (1998): Compared language ability of children between 1 and 3 years of age.

There are advantages to early acquisition

Language deprivation in deaf children

More than just a scientific question...

Language creation

Language creation

Deaf children with no prior linguistic experience

New school for the deaf (Managua, 1997)

Desire for social interaction and communication

Language creation

Deaf children with no prior linguistic experience

New school for the deaf (Managua, 1997)

Desire for social interaction and communication
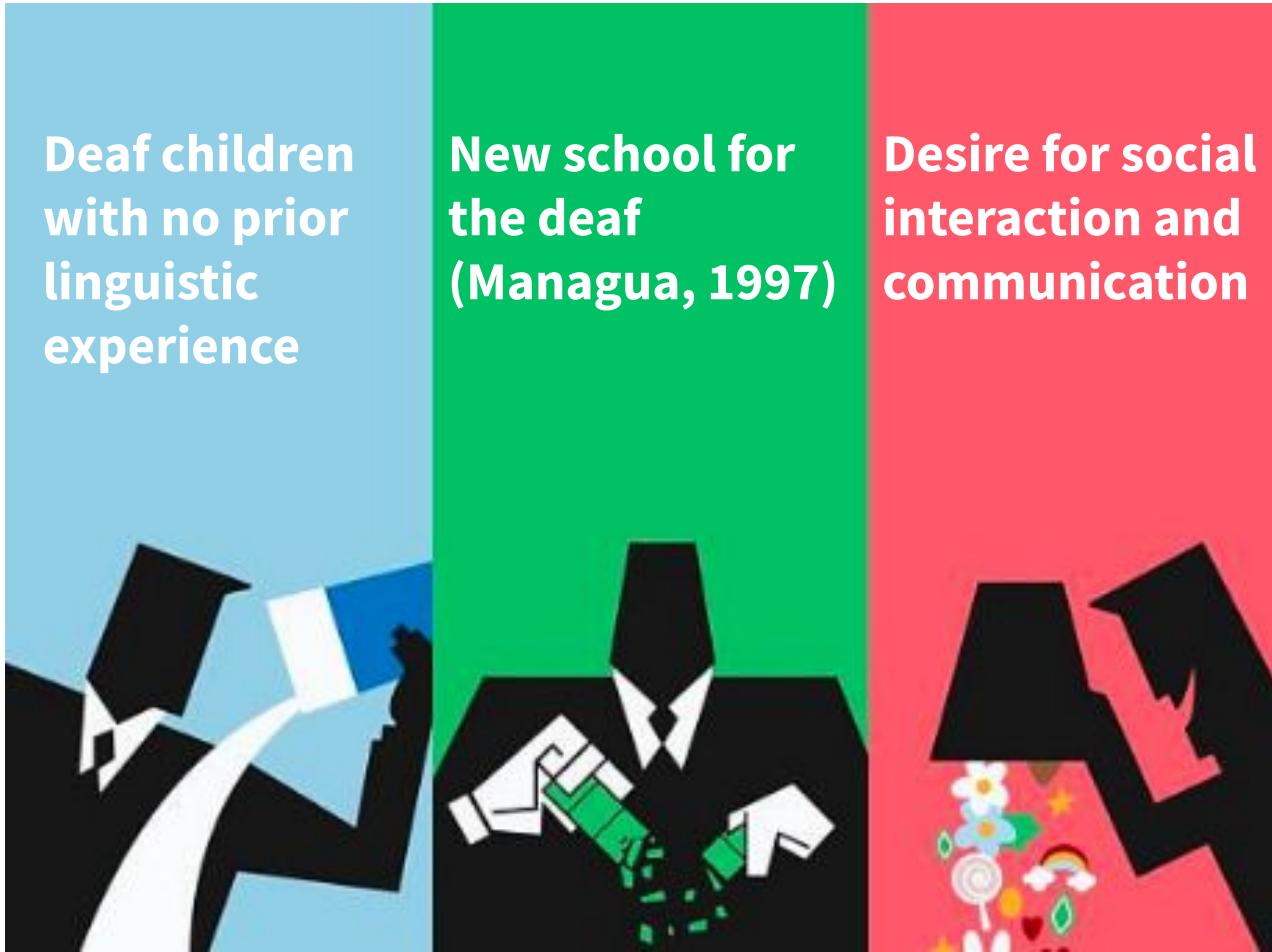
Nicaraguan Sign Language!

Language creation

Deaf children with no prior linguistic experience

New school for the deaf (Managua, 1997)

Desire for social interaction and communication

Nicaraguan Sign Language!

Does this have structure that resembles other languages?

Language creation

Without spatial modulation
(action performed by noun
signed before)

With spatial modulation
(action is performed by noun
placed in the spatial location)

"see"

"pay"

Language creation

Without spatial modulation
(action performed by noun signed before)

With spatial modulation
(action is performed by noun placed in the spatial location)

"see"

"pay"

Language creation

Senghas & Coppola (2001)

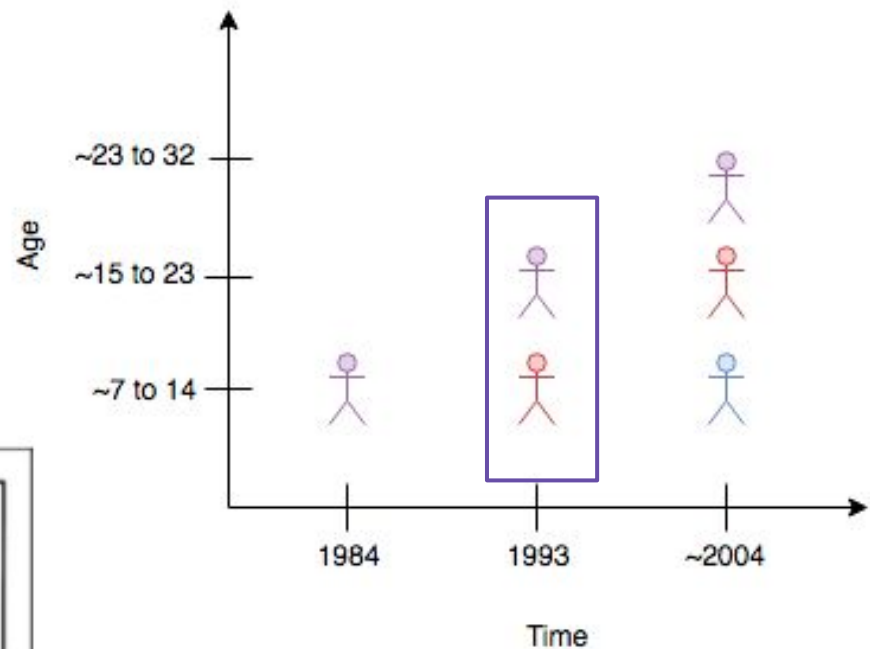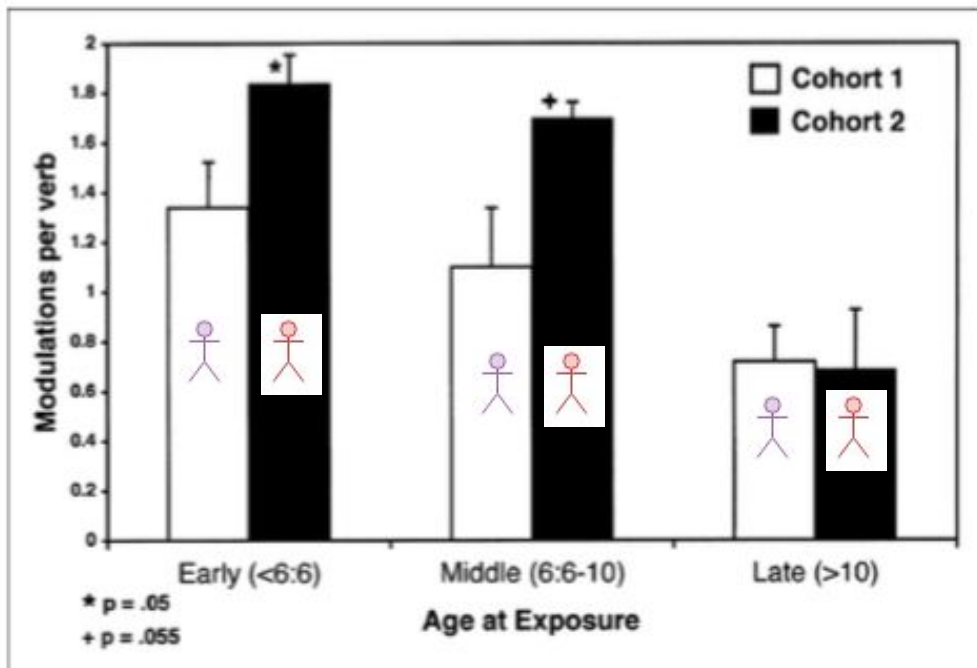# Language creation

Senghas & Coppola (2001)

Fig. 2. Mean number of spatial modulations per verb produced by early-, middle-, and late-exposed signers of the first and second cohorts. Asterisks indicate a significant difference between cohorts within an age-at-exposure group.

Language creation

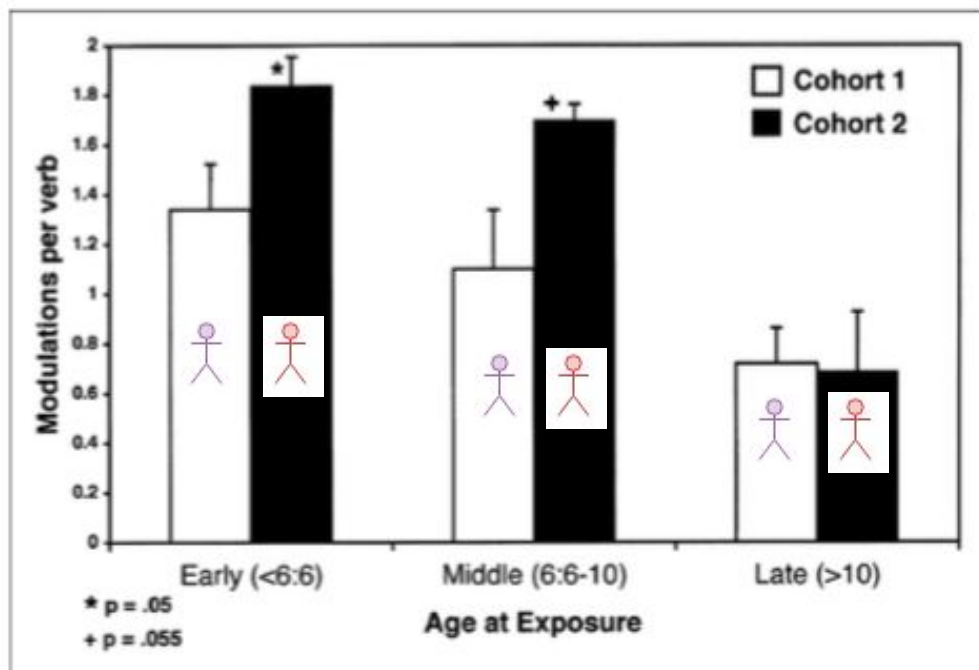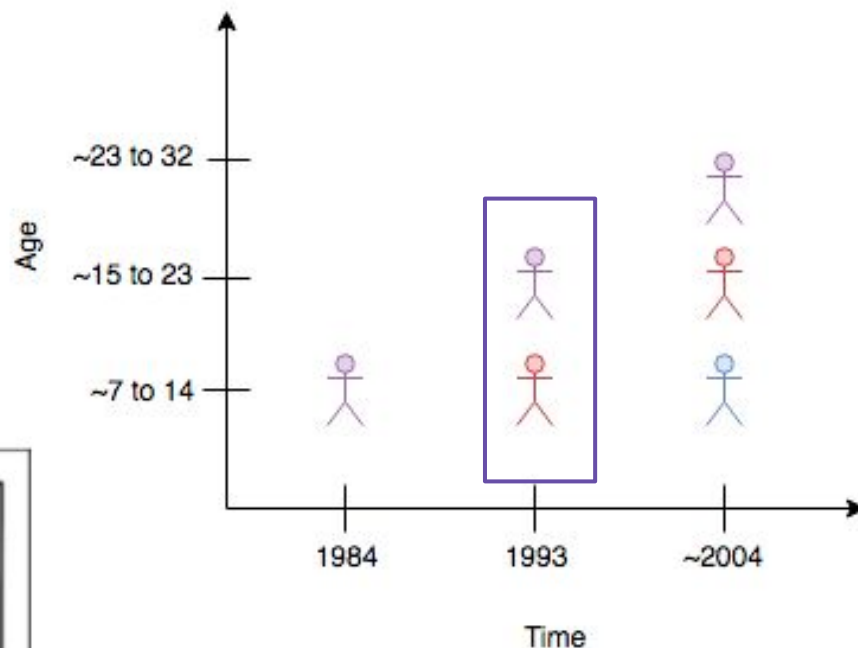Senghas & Coppola (2001)

Fig. 2. Mean number of spatial modulations per verb produced by early-, middle-, and late-exposed signers of the first and second cohorts. Asterisks indicate a significant difference between cohorts within an age-at-exposure group.

Cohort 2 has more spatial modulation than cohort 1

Spatial modulation decreases as age at exposure increases

Language creation

One action for both manner and path

Distinct actions for manner and path (performed sequentially)

Language creation

One action for both manner and path
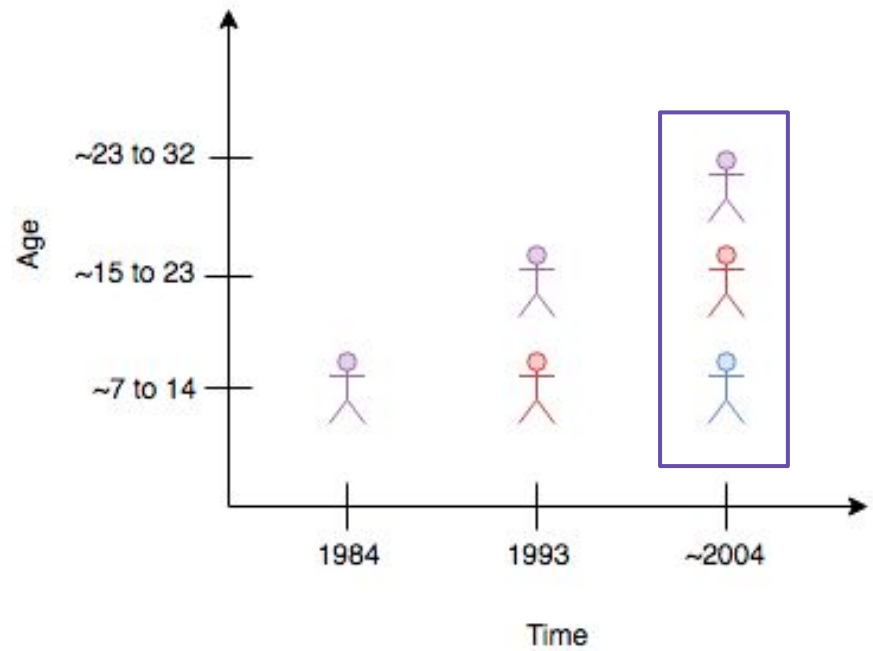

Distinct actions for manner and path (performed sequentially)

More complex/ abstract

Language creation

Senghas et al (2004)
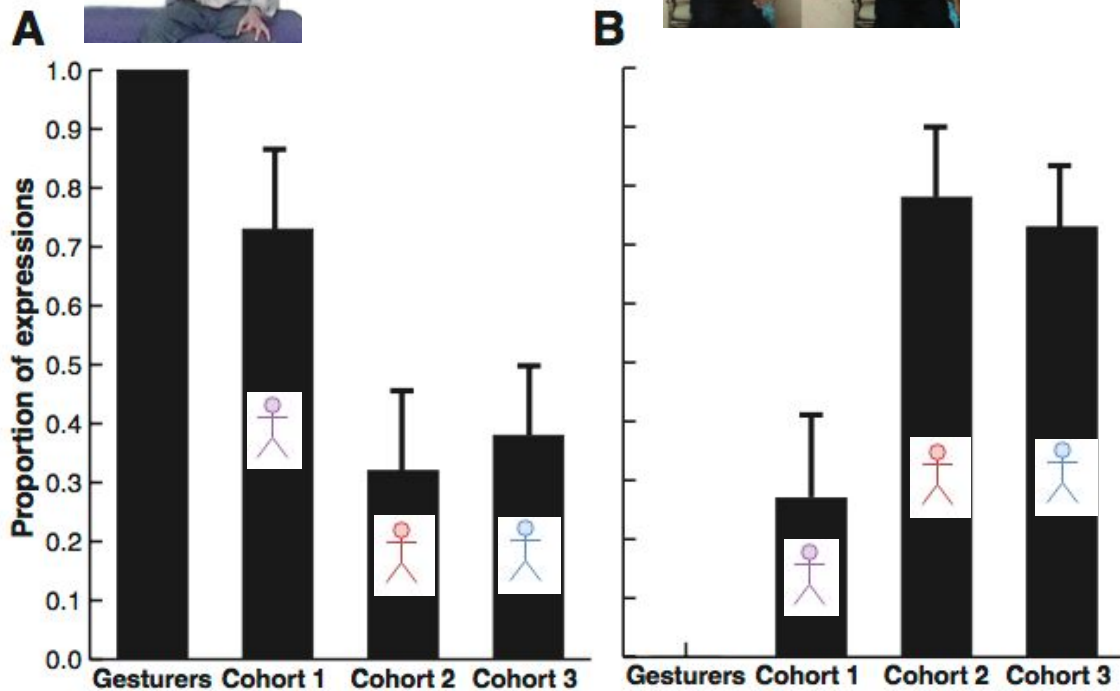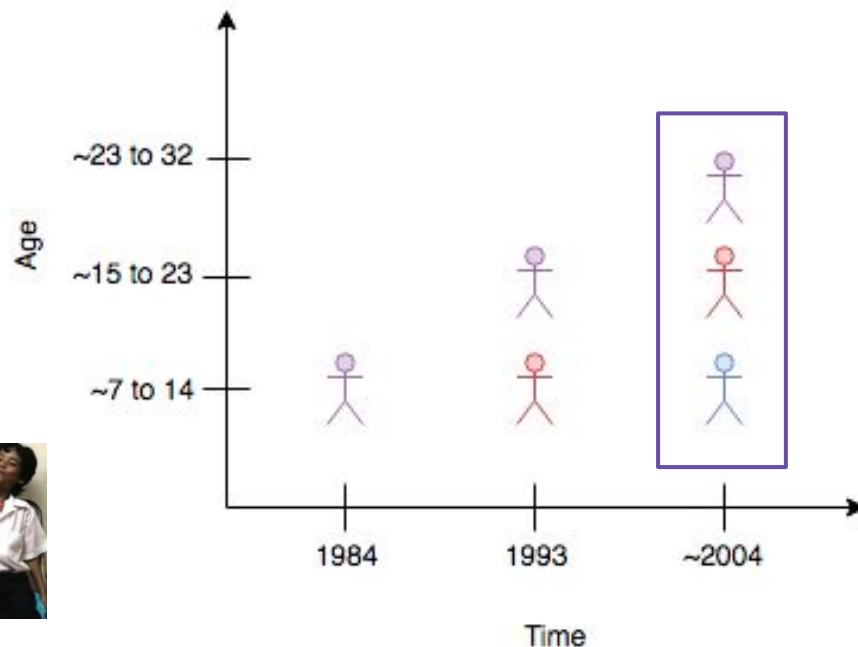
Language creation

Senghas et al (2004)

Cohort 2 and 3 use more the distinct action for manner and event

Language creation

Conclusions

- A group of children with no linguistic experience created a language that has structure resembling other signed/ spoken languages

- This language evolved over time with each incoming cohort — so children were responsible for the change.

Language creation

Conclusions

- A group of children with no linguistic experience created a language that has structure resembling other signed/ spoken languages

- This language evolved over time with each incoming cohort — so children were responsible for the chan

Language creation

Conclusions

- A group of children with no linguistic experience created a language that has structure resembling other signed/ spoken languages

- This language evolved over time with each incoming cohort — so children were responsible for the chan

Senghas:

https://www.youtube.com/watch?v=FTPGmKoDk0Y

Syntactic bootstrapping

Glietman (1990)
- Meaning of verbs cannot be acquired from observation alone
  - Many verbs are identical except for the perspective (e.g. chasing vs fleeing)
    - The hounds are chasing the fox.
    - The fox is fleeing from the hounds
  - Many verbs don't refer to states that are observable.

Syntactic bootstrapping

Glietman (1990)
- Meaning of verbs cannot be acquired from observation alone
    - Many verbs are identical except for the perspective (e.g. chasing vs fleeing)
        - The hounds are chasing the fox.
        - The fox is fleeing from the hounds
    - Many verbs don't refer to states that are observable.

- Children use syntax to infer the meaning of the verb. (bootstrapping)

- The link between syntax and meaning is likely innate
    - Where in the sentence the subject occurs (UG)
    - The subject is central to the event

Syntactic bootstrapping

Naigles (1990)

- Children between 1;1 and 2;3

- Preferential looking paradigm

  - Two simultaneous videos with a sentence presented in the background
  - Consistent video, longer looking time
    - Note: Different from habituation studies where surprising = longer looking time.
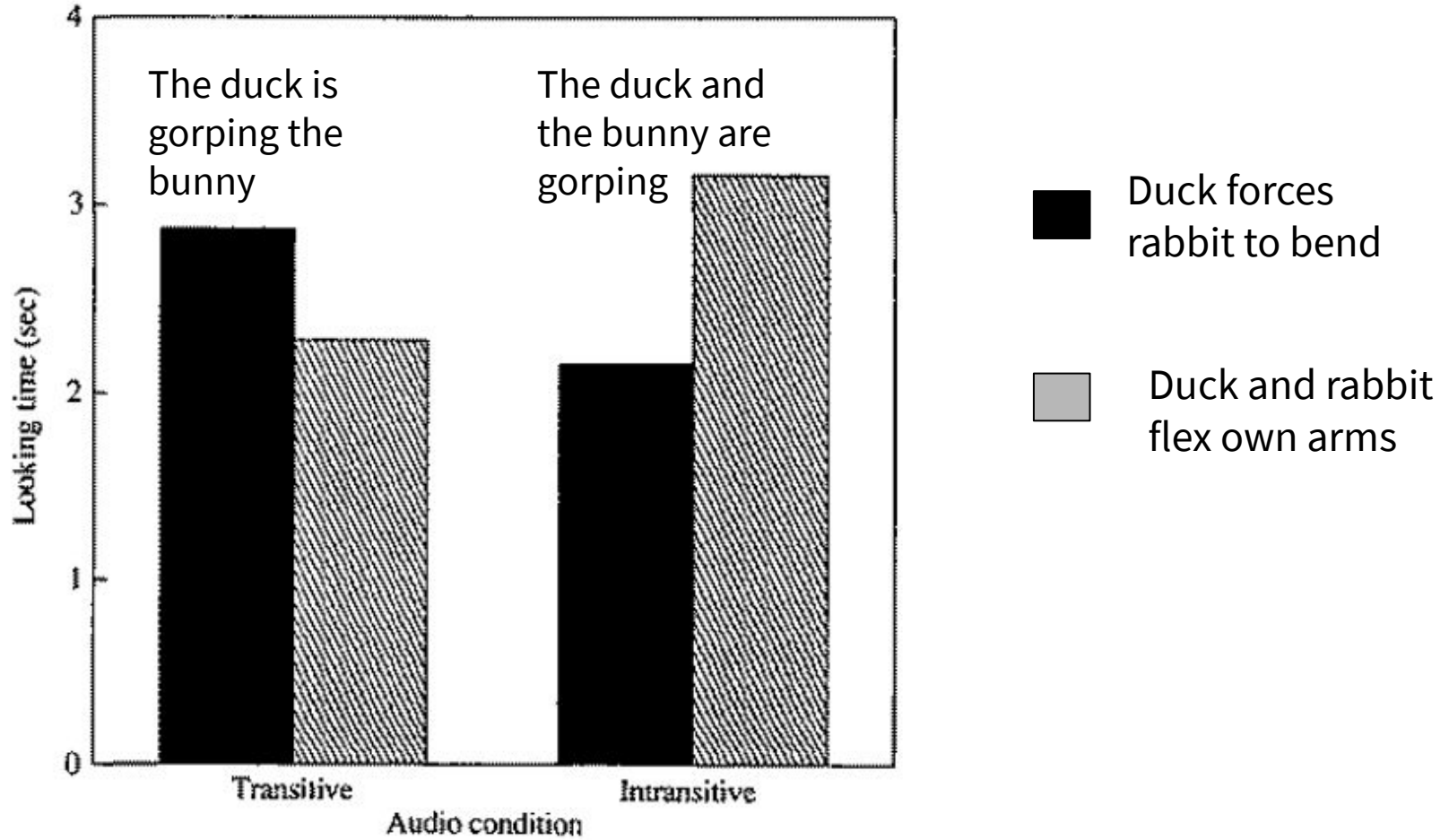
Syntactic bootstrapping

Fig. 3. Mean visual fixation during the test trials to the causative and non-causative actions, for the Transitive and Intransitive Audio conditions. ■, Causative action; □, non-causative action.

Outline

Arguments for innate, domain specific and modular mechanism

- Chomsky's argument (UG and LAD)
- Critical period
  - Genie
  - Language deprivation in deaf children
- Language creation (Nicaraguan Sign Language)
- Syntactic bootstrapping

Arguments for experience dependent and domain general mechanism

- Criticism of Chomsky's argument
- Statistical learning in word segmentation
- Role of bias in word learning
- Sensitivity to statistical regularities in syntactic structure
- (Morpho)-Syntax in connectionist networks

Arguments for an experience dependent and domain general mechanism

Criticisms of Chomsky's argument

- Poverty of the stimulus argument is an empirical claim, not a logical argument that is necessarily true
  - Principles and parameters do not have to be innate
  - Modular system like LAD not required

- What else can be innate?

  - Learning mechanisms
    - E.g. Rely on statistical regularities in the environment to make predictions about what is most likely to occur next
  - Inductive biases
    - E.g. Automatically pay attention to shapes (remember preference for faces?)
    - E.g. Have a preference for hierarchical structure — i.e. things/events/ideas are made of smaller parts that can be combined

# Criticisms of Chomsky's argument

- Poverty of the stimulus argument is an empirical claim, not a logical argument that is necessarily true
    - Principles and parameters do not have to be innate
    - Modular system like LAD not required

- What else can be innate?   **These can be domain general**

    - Learning mechanisms
        - E.g. Rely on statistical regularities in the environment to make predictions about what is most likely to occur next
    - Inductive biases
        - E.g. Automatically pay attention to shapes (remember preference for faces?)
        - E.g. Have a preference for hierarchical structure — i.e. things/events/ideas are made of smaller parts that can be combined

Statistical learning in word segmentation

How do we segment continuous flow of speech into words?

can   you   pass   me the   salt
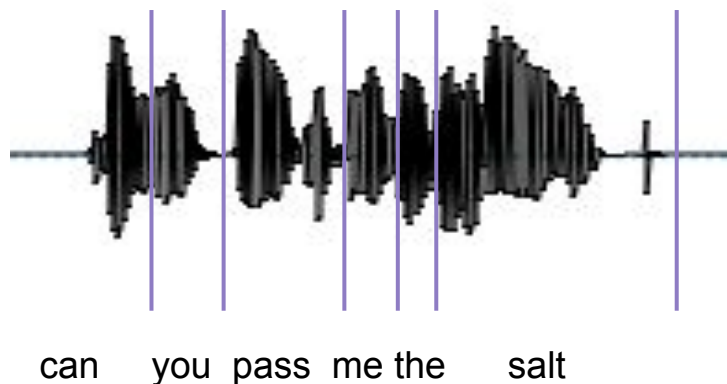
Possible strategy: Rely on transitional probabilities

$p(/s/ \mid /pa/) > p(/mɪ/ \mid /s/)$

Mark a word boundary when the transitional probabilities are the lowest

Maximum likelihood estimate of transitional probabilities

$$p\,(a \mid b) = \frac{count(a,b)}{count(a)}$$

Statistical learning in word segmentation

Saffran et al (1996)

Three "words" presented back to back. Each word could be followed by either of the two words or itself

- bidaku   p(bi | ku) = p(pa | ku) = p(go | ku) = ⅓ = 0.33
- padoti   p(bi | ti) = p(pa | ti) = p(go | ti) = ⅓ = 0.33
- golabu   p(bi | bu) = p(pa | bu) = p(go | bu) = ⅓ = 0.33

All other transitional probabilities are 1.

Experiment 1: Infants surprised by 'kudabi' (impossible sequence)

Experiment 2: Infants surprised by 'kupado' (possible non-word sequence)

Statistical learning in word segmentation

Saffran et al (1996)

Three "words" presented back to back. Each word could be followed by either of the two words or itself

- bidaku $p(bi \mid ku) = p(pa \mid ku) = p(go \mid ku) = ⅓ = 0.33$
- padoti $p(bi \mid ti) = p(pa \mid ti) = p(go \mid ti) = ⅓ = 0.33$
- golabu $p(bi \mid bu) = p(pa \mid bu) = p(go \mid bu) = ⅓ = 0.33$

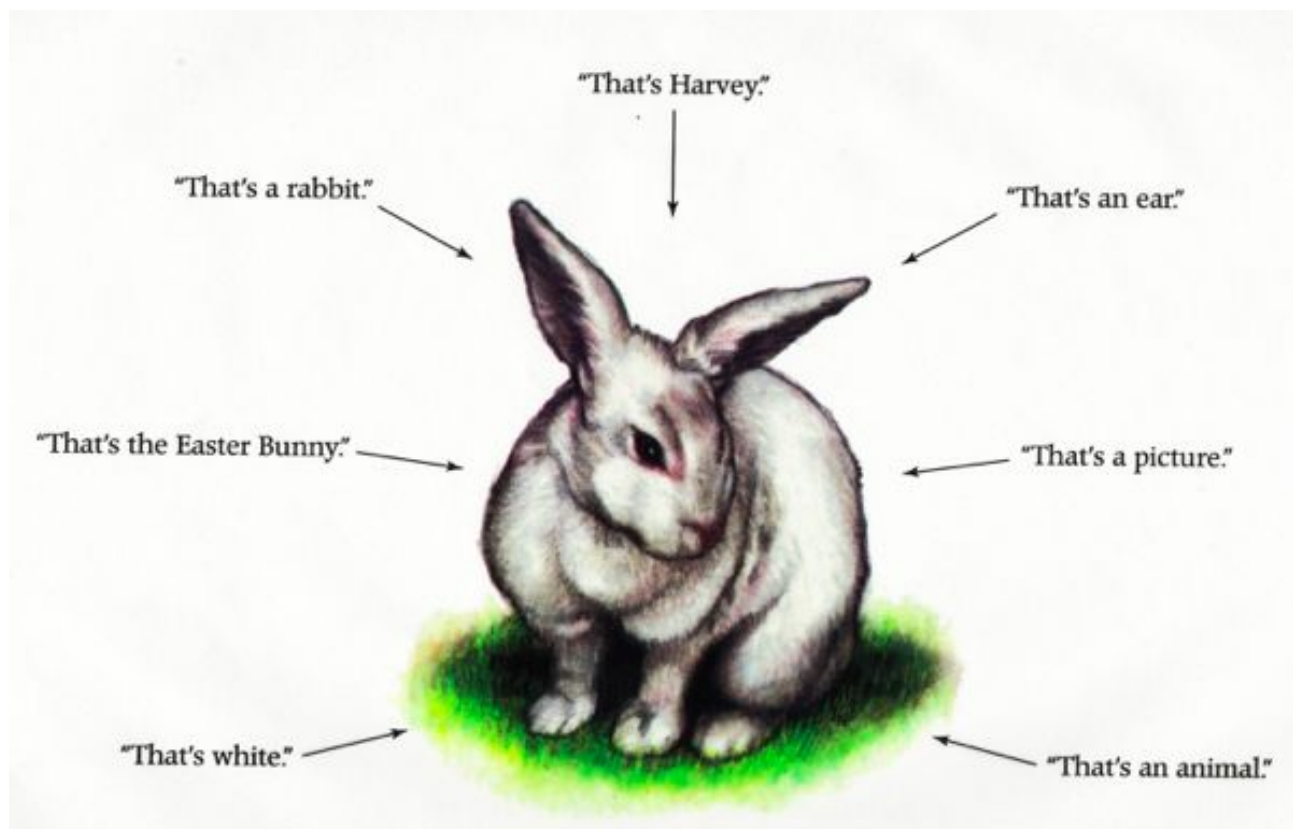All other transitional probabilities are 1.

Experiment 1: Infants surprised by 'kudabi' (impossible sequence)

Experiment 2: Infants surprised by 'kupado' (possible non-word sequence)

Innate learning mechanism?

Role of bias in word learning

Quine's problem: How do we learn what words mean given that different descriptions can be used to talk about the same percept?
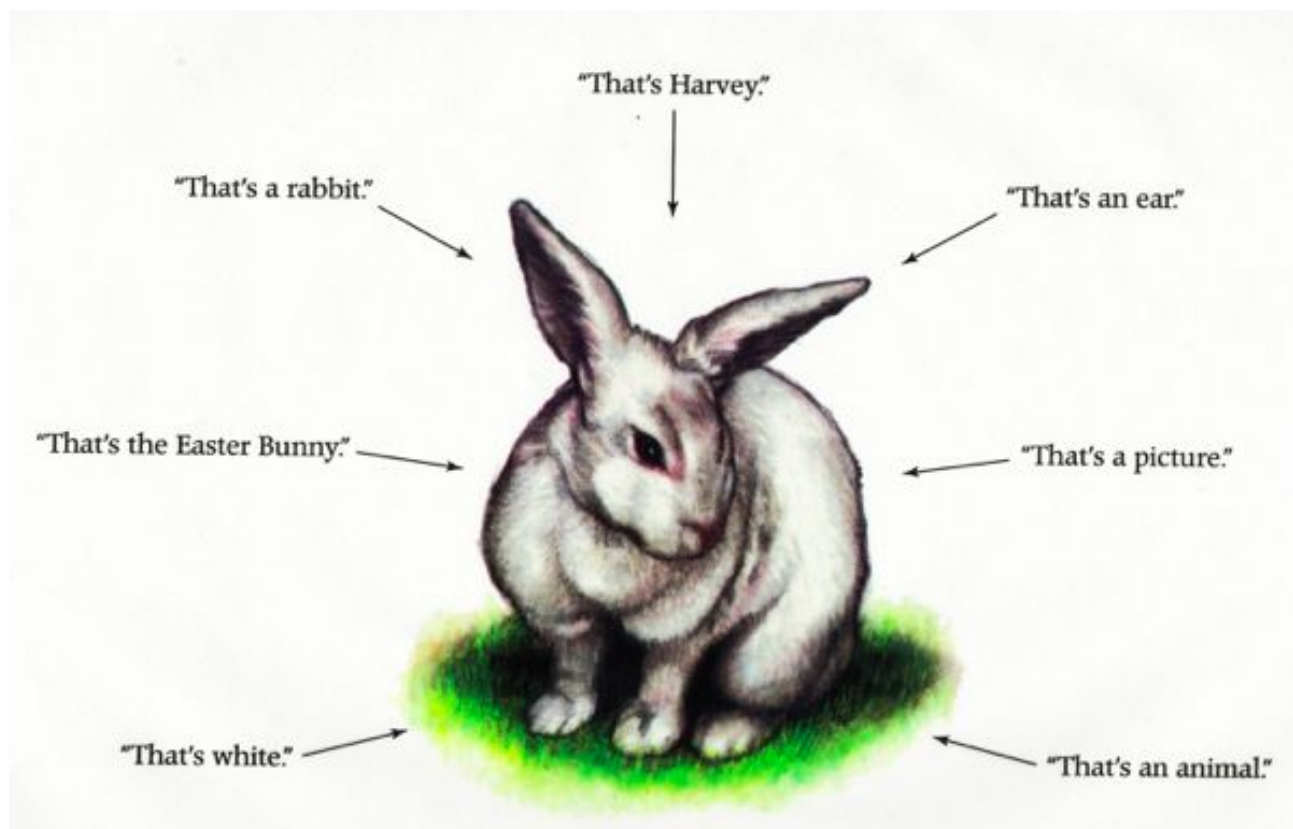
Role of bias in word learning

Quine's problem: How do we learn what words mean given that different descriptions can be used to talk about the same percept?

"That's Harvey."

"That's a rabbit."

"That's an ear."

"That's the Easter Bunny."

"That's a picture."

"That's white."

"That's an animal."

Possible explanation: Inductive bias

Role of bias in word learning

Markman & Wachtel (1988) — rephrased in terms of bias

- **Whole object bias:** People tend to talk about whole objects. So labels are more likely to refer to whole objects
  - "This is a trachea" — child selects the whole object
- **Taxonomic constraint**: New labels (in parallel structure) are more likely to relate to taxonomically related object (dog-cat) than thematically related object (dog-bone)
  - "See this fep ? (pointing to a dog). Find another" — child selects cat
  - "See this? (pointing to a dog). Find another" — child selects bone

Landau et al (2001)

- **Shape bias:** When shown a novel named object, they tend to judge objects with the same shape as belonging to same category

But these are examples of word learning. Wasn't Chomsky's argument about syntax? Can you get syntax acquisition from domain general biases or learning mechanisms?

# Sensitivity to statistical regularities in syntactic structure

- Children are sensitive to statistical regularities in syntactic structure — i.e. p(structure | previous context)
- "I saw the house with _____ "
  - a garden : NP attachment
  - my friend: VP attachment
- When in an experimental environment with more NP attachments, children around the age of 5 and 6 will begin to pick more NPs (Havron et al, 2018)
- Adults also display widespread effects of statistical regularities — i.e. they are more surprised when they see an improbable structure

# Sensitivity to statistical regularities in syntactic structure

- Children are sensitive to statistical regularities in syntactic structure — i.e. p(structure | previous context)
- "I saw the house with _____ "
  - a garden : NP attachment
  - my friend: VP attachment
- When in an experimental environment with more NP attachments, children around the age of 5 and 6 will begin to pick more NPs (Havron et al, 2018)
- Adults also display widespread effects of statistical regularities — i.e. they are more surprised when they see an improbable structure

No well fleshed out theory about language acquisition in terms of just sensitivity to statistical regularities

(Morpho-)Syntax in connectionist networks
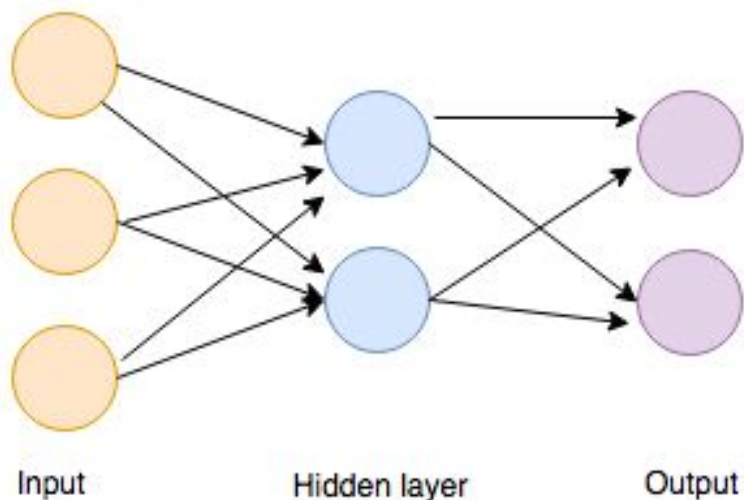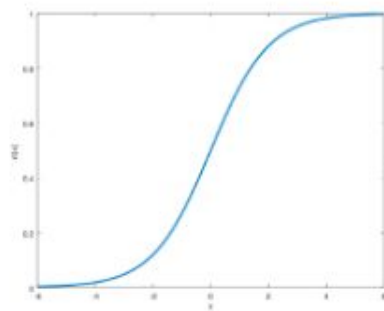
Input   Hidden layer   Output
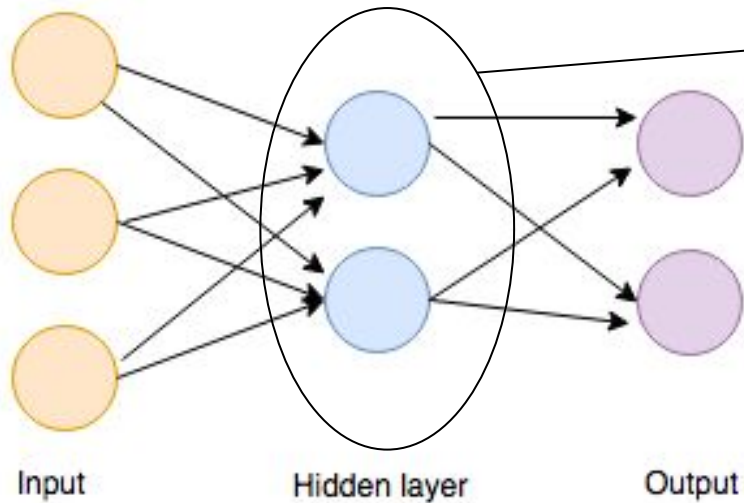
Activation of input nodes *
weight

Some differentiable
function

Activation of current
nodes

(Morpho-)Syntax in connectionist networks

Tends to be really complex architectures with thousands of nodes

Input

Hidden layer

Output
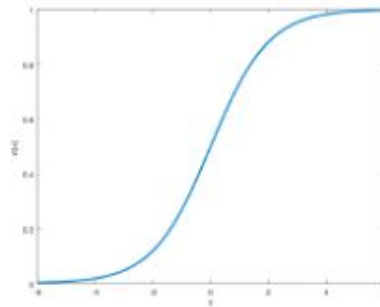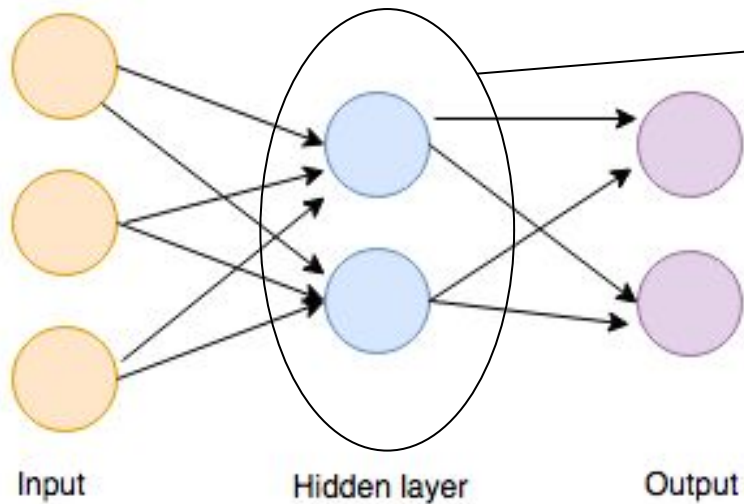
Activation of input nodes * weight

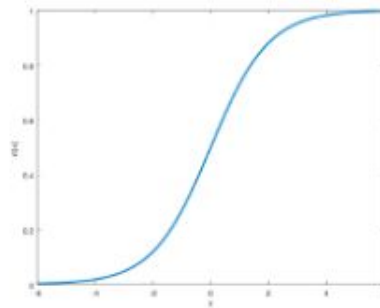Some differentiable function

Activation of current nodes

(Morpho-)Syntax in connectionist networks

Tends to be really complex architectures with thousands of nodes

Essentially just general purpose pattern detectors

Input

Hidden layer

Output

Activation of input nodes * weight

Some differentiable function

Activation of current nodes

(Morpho-)Syntax in connectionist networks

McLelland & Rumelhart (1986)

- Connectionist network that acquired past tense morphology

- Made similar errors as children with over generalization initially and then acquiring it

- No explicit rules were encoded, but the network still behaved as if it were following rules — i.e. rules emerged

More recent work: Networks exhibit behaviours that suggest they represent (to a certain extent) hierarchical syntactic dependencies

(Morpho-)Syntax in connectionist networks

McLelland & Rumelhart (1986)

- Connectionist network that acquired past tense morphology

- Made similar errors as children with over generalization initially and then acquiring it

- No explicit rules were encoded, but the network still behaved as if it were following rules — i.e. rules emerged

More recent work: Networks exhibit behaviours that suggest they represent (to a certain extent) hierarchical syntactic dependencies

Networks need a lot of training data. Additionally, even state of the art networks fail on difficult sentences

Summary / Things to think about

Arguments for innate, domain specific and modular mechanism

- Chomsky's argument (UG and LAD)
- Critical period
  - Genie
  - Language deprivation in deaf children
- Language creation (Nicaraguan Sign Language)
- Syntactic bootstrapping

Arguments for experience dependent and domain general mechanism

- Criticism of Chomsky's argument
- Statistical learning in word segmentation
- Role of bias in word learning
- Sensitivity to statistical regularities in syntactic structure
- (Morpho)-Syntax in connectionist networks

Summary / Things to think about

Arguments for innate, domain specific and modular mechanism

- Chomsky's argument (UG and LAD)
- Critical period
  - Genie
  - Language deprivation in deaf children
- Language creation (Nicaraguan Sign Language)
- Syntactic bootstrapping

Not necessarily true. Even if it is true, is LAD a sufficient explanation?

Can these empirical phenomena be explained without LAD?

Arguments for experience dependent and domain general mechanism

- Criticism of Chomsky's argument
- Statistical learning in word segmentation
- Role of bias in word learning
- Sensitivity to statistical regularities in syntactic structure
- (Morpho)-Syntax in connectionist networks

More parsimonious/ neurally plausible?

More work needs to be done to explain how these mechanisms can lead to complex linguistic behaviour

Fun https://www.youtube.com/watch?v=1KUP4p1cGsk